ICS 35.240
CCS L00/09

团　　体　　标　　准

T/ISC XXX—XXXX

数字安全大模型应用成熟度模型与评估

Digital Security Large Language Model Application Maturity Model （DSL^2M^2） and Assessment

在提交反馈意见时，请将您知道的相关专利与支持性文件一并附上。

征求意见稿

XXXX－XX－XX 发布　　XXXX－XX－XX 实施

中 国 互 联 网 协 会　　发布

目　　次

前　言

本文件按照GB/T 1.1—2020《标准化工作导则　第1部分：标准化文件的结构和起草规则》的规定起草。

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别专利的责任。

本文件由中国互联网协会提出并归口。

本文件起草单位：中国信息通信研究院、南方电网互联网服务有限公司、国网冀北电力有限公司智能配电网中心(北戴河供电保障指挥中心)、大连银行股份有限公司、北京安普诺信息技术有限公司、闪捷信息科技有限公司、亚信安全科技股份有限公司、云南易数科技有限责任公司、北京宇卫科技有限公司、北京弘畅教育科技有限公司、北京奕华科技有限公司

本文件主要起草人：马英轩、侯洪磊、陈丽娜、桂媛、夏武、张润学、陈志敏、张子超、张涛、王越、苏伟华、杨婷、赵吕、苏丕云、王霖岩、胡珏。

引　言

为推动落实《全球人工智能治理倡议》，打造可审核、可监督、可追溯、可信赖的人工智能技术。同时为了组织能够更有效、更安全的使用人工智能技术，由中国信通院牵头，联合国内具备自主研发能力的网络安全厂商以及将人工智能技术应用到日常网络安全运营工作中的企业，共同编制本文件。数字安全大模型应用成熟度模型（Digital Security Large Language Model Application Maturity Model，DSL^2M^2）是面向数字安全领域构建的大模型应用能力评价框架。DSL^2M^2借鉴能力成熟度模型（Capability Maturity Model，CMM）和能力成熟度模型集成（Capability Maturity Model Integration，CMMI）思想，以“能力建设、业务融合、安全治理、持续优化”为主线，构建覆盖产品基础能力、模型工程能力、业务融合能力、安全治理能力、组织运营能力和持续优化能力的成熟度模型体系。DSL^2M^2适用于数字安全产品提供方、数字安全运营组织及第三方评估机构开展能力建设、成熟度评估和认证工作。通过建立统一的成熟度等级体系、能力域体系和评估方法，引导组织实现数字安全大模型从单点试用向体系化应用、从辅助决策向智能协同、从局部优化向持续进化的发展演进。

数字安全大模型应用能力成熟度评估规范

1 范围

本文件规定了数字安全大模型应用成熟度模型的总体框架、成熟度等级、能力域体系、成熟度要求和评估方法。

本文件构建了覆盖产品基础能力、模型工程能力、业务融合能力、安全治理能力、组织运营能力和持续优化能力的数字安全大模型应用成熟度模型（Digital Security Large Language Model Application Maturity Model，DSL^2M^2），用于评价数字安全领域大模型应用的能力建设水平、应用成熟程度和持续发展能力。

本文件适用于：

1. 数字安全大模型产品、平台和解决方案提供方开展能力建设、能力评估和产品认证；
2. 政府部门、企事业单位、安全运营机构及其他组织开展数字安全大模型应用能力建设、成熟度评估和管理改进；
3. 第三方评估机构开展数字安全大模型应用成熟度评估、评价和咨询服务；
4. 行业组织、科研机构和相关单位开展数字安全大模型应用能力研究、能力评价和标准化工作。

本文件适用于数字安全领域大模型、智能体及相关智能化应用的成熟度评估，不适用于基础模型算法性能测试、通用人工智能能力测试和专项安全测评。

本文件规定的成熟度模型可用于产品型评估（DSL^2M^2-P）和组织型评估（DSL^2M^2-O）。

本文件不对大模型的技术路线、模型架构、参数规模、部署模式和实现方式作限定，评估对象可根据自身业务需求和应用场景选择适用技术方案。

2 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中，注日期的引用文件，仅该日期对应的版本适用于本文件；不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

GB/T 19000—2016 质量管理体系 基础和术语

GB/T 25000.10—2016 系统与软件工程 系统与软件质量要求和评价（SQuaRE） 第10部分：系统与软件质量模型

GB/T 25000.51—2016 系统与软件工程 系统与软件质量要求和评价（SQuaRE） 第51部分：就绪可用软件产品（RUSP）的质量要求和测试细则

GB/T 25069—2022 信息安全技术 术语

GB/T 29246—2017 信息技术 安全技术 信息安全管理体系 概述和词汇

GB/T 20261—2020 信息安全技术 系统安全工程 能力成熟度模型

GB/T 35273 信息安全技术 个人信息安全规范

GB/T 37988 信息安全技术 数据安全能力成熟度模型

GB/T 41867 信息技术 人工智能 术语

GB/T 42576 信息技术 人工智能 管理体系

GB/T 43439 数字化转型成熟度模型与评估
GB/T 45989 软件过程能力成熟度模型
ISO/IEC 42001 Information technology — Artificial intelligence — Management system
ISO/IEC 22989 Information technology — Artificial intelligence concepts and terminology
ISO/IEC 23894 Information technology — Artificial intelligence — Guidance on risk management

3 术语和定义

GB/T 25069、GB/T 29246、GB/T 41867界定的以及下列术语和定义适用于本文件。

3.1 大模型 Large Language Model

基于海量数据训练形成、具有大规模参数和通用任务处理能力的人工智能模型。
注：大模型通常具备自然语言理解、内容生成、知识推理、任务规划和工具调用等能力。

3.2

3.2 数字安全大模型 Digital Security Large Language Model

面向数字安全领域构建或优化的大模型，具备安全知识理解、安全分析研判、安全运营支撑和安全决策辅助等能力。
注：数字安全大模型可基于通用大模型构建，也可采用行业专用模型构建。

3.3 智能体 Agent

能够感知环境、理解目标、自主规划并执行任务的智能实体。
注：智能体可基于大模型构建，并通过调用工具、系统或其他智能体完成复杂任务。

3.4 数字员工 Digital Employee

在组织中承担特定岗位职责，通过大模型、智能体等技术实现任务执行、业务协同和决策支持的数字化智能主体。
注：数字员工可独立执行部分任务，也可与人类员工协同完成工作。

3.5 数字安全大模型应用成熟度模型 Digital Security Large Language Model Application Maturity Model（DSL^2M^2）

用于评价组织或产品在数字安全领域应用大模型能力建设水平、应用成熟程度和持续发展能力的模型体系。
注：DSL^2M^2由成熟度等级、能力域、成熟度要求和评估方法组成。

3.6 成熟度等级 Maturity Level

反映评估对象能力建设水平和发展阶段的等级划分。
注：本文件将成熟度等级划分为L1初始级、L2可管理级、L3已定义级、L4量化管理级和L5优化创新级。

3.7 能力域 Capability Domain

数字安全大模型应用过程中需要具备的关键能力集合。

注：本文件包括产品基础能力、模型工程能力、业务融合能力、安全治理能力、组织运营能力和持续优化能力六个能力域。

3.8 能力子域 Capability Sub-domain

能力域下进一步细分的具体能力单元，用于描述特定能力方向的建设要求和成熟度特征。

3.9 产品型评估 Product Assessment

针对数字安全大模型产品、平台或解决方案开展的成熟度评估活动。

注：产品型评估结果表示为DSL^2M^2-P。

3.10 组织型评估 Organization Assessment

针对政府部门、企事业单位、安全运营机构等组织开展的成熟度评估活动。

注：组织型评估结果表示为DSL^2M^2-O。

3.11 人机协同 Human-AI Collaboration

人类员工与数字员工、智能体共同参与业务活动并协同完成任务的工作模式。

3.12 能力运营 Capability Operation

对数字安全大模型、数字员工和智能体能力进行持续监测、评价、优化和管理的活动。

3.13 持续优化 Continuous Improvement

基于评估结果、运行数据和业务反馈，对数字安全大模型应用能力进行持续改进和能力提升的过程。

3.14 安全治理 Security Governance

围绕数字安全大模型应用建立的制度、流程、技术和组织管理体系，用于实现风险控制、合规管理和可信运行。

3.15 能体治理 Agent Governance

针对智能体和数字员工建立身份管理、权限管理、行为管理、审计追溯和风险控制机制的治理活动。

4 符号和缩略语

下列缩略语适用于本文件。

AI：Artificial Intelligence（人工智能）

CMM：Capability Maturity Model（能力成熟度模型）

CMMI：Capability Maturity Model Integration（能力成熟度模型集成）

DSL^2M^2：Digital Security Large Language Model Application Maturity Model（数字安全大模型应用成熟度模型）

DSL^2M^2-P：Digital Security Large Language Model Application Maturity Model for Product（产品型成熟度评估）

DSL^2M^2-O：Digital Security Large Language Model Application Maturity Model for Organization（组织型成熟度评估）

LLM：Large Language Model（大语言模型）
RAG：Retrieval-Augmented Generation（检索增强生成）
MCP：Model Context Protocol（模型上下文协议）
SOC：Security Operations Center（安全运营中心）
TTFT：Time To First Token（首Token响应时间）

5 数字安全大模型应用成熟度模型

5.1 模型框架

5.1.1 整体框架

数字安全大模型应用成熟度模型（Digital Security Large Language Model Application Maturity Model，DSL^2M^2）用于评价组织在数字安全领域应用大模型技术的能力建设水平、业务融合水平、安全治理水平和持续优化能力。DSL^2M^2借鉴能力成熟度模型（Capability Maturity Model，CMM）和能力成熟度模型集成（Capability Maturity Model Integration，CMMI）的理念，以能力建设、业务融合、安全治理和持续优化为主线，构建覆盖产品基础能力、模型工程能力、业务融合能力、安全治理能力、组织运营能力和持续优化能力的成熟度模型体系。

DSL^2M^2由成熟度等级、能力域和特征要求矩阵三部分构成。

5.1.2 成熟度等级

成熟度等级用于描述数字安全大模型应用能力的发展阶段、能力水平和演进目标。

成熟度等级按照组织数字安全大模型应用能力的发展水平划分为五个等级，各等级由低到高逐级递进。

5.1.3 能力域

能力域是组织开展数字安全大模型应用过程中所需具备的关键能力集合。

能力域用于描述组织在技术能力建设、业务应用、安全治理和运营管理等方面应达到的能力要求。

5.1.4 特征要求矩阵

特征要求矩阵用于描述各能力域在不同成熟度等级下应具备的能力特征和要求。

组织应满足对应成熟度等级规定的能力要求，并持续提升各能力域能力水平。

5.2 成熟度等级

5.2.1 等级划分

数字安全大模型应用成熟度划分为五个等级。

各等级由低到高依次为：

a）L1 初始级（Initial）；
b）L2 可管理级（Managed）；
c）L3 已定义级（Defined）；
d）L4 量化管理级（Quantitatively Managed）；
e）L5 优化创新级（Optimizing）。

5.2.2 L1 初始级

组织开始关注数字安全大模型应用，并在部分场景开展探索与实践。

数字安全大模型应用主要依赖局部项目或个体经验驱动，尚未形成统一规划、标准流程和管理机制。

5.2.3 L2 可管理级

组织已建立数字安全大模型应用的基本管理机制，能够对相关项目进行规划、建设和管理。

数字安全大模型已应用于部分业务场景，并建立基础制度规范和运行管理机制。

5.2.4 L3 已定义级

组织已形成统一的数字安全大模型应用体系和标准化流程。

数字安全大模型已融入核心业务活动，形成组织级能力建设和应用推广机制。

5.2.5 L4 量化管理级

组织建立数字安全大模型应用量化管理体系。

能够对模型能力、业务价值、安全风险和运营效果进行持续监测、量化分析和评估，并基于数据开展持续改进。

5.2.6 L5 优化创新级

组织实现数字安全大模型应用的持续优化和创新发展。

数字安全大模型成为组织数字安全能力体系的重要组成部分，并具备持续演进、智能协同和创新应用能力。

5.3 能力域

5.3.1 概述

DSL^2M^2从产品能力、工程能力、业务能力、治理能力和运营能力等方面构建能力域体系。

能力域包括：

a）产品基础能力；

b）模型工程能力；

c）业务融合能力；

d）安全治理能力；

e）组织运营能力；

f）持续优化能力。

5.3.2 产品基础能力

产品基础能力用于评价数字安全大模型产品或系统自身所具备的基础技术能力。

产品基础能力主要包括模型能力、性能能力、资源利用能力和专业能力等方面。

5.3.3 模型工程能力

模型工程能力用于评价数字安全大模型在部署实施、集成应用、能力扩展和工程化支撑等方面的能力水平。

模型工程能力主要包括部署能力、知识增强能力、工具协同能力和系统集成能力等方面。

5.3.4 业务融合能力

业务融合能力用于评价数字安全大模型与数字安全业务场景融合应用的深度和广度。

业务融合能力主要包括业务支撑能力、流程融合能力和智能协同能力等方面。

5.3.5 安全治理能力

安全治理能力用于评价数字安全大模型全生命周期安全管理与风险控制能力。

安全治理能力主要包括数据安全治理、模型安全治理、智能体安全治理、合规治理和安全运营等方面。

5.3.6 组织运营能力

组织运营能力用于评价组织开展数字安全大模型应用过程中形成的管理、运营和保障能力。

组织运营能力主要包括战略规划、组织机制、身份管理、人机协同和人才建设等方面。

5.3.7 持续优化能力

持续优化能力用于评价组织开展数字安全大模型持续改进和创新发展的能力水平。

持续优化能力主要包括能力评价、能力运营、知识运营、模型演进和创新发展等方面。

5.4 特征要求矩阵

5.4.1 概述

特征要求矩阵用于描述各能力域在不同成熟度等级下应达到的能力要求。

组织成熟度等级应根据各能力域能力水平综合确定。

5.4.2 能力域与成熟度等级对应关系

数字安全大模型应用成熟度模型能力域与成熟度等级对应关系见表1。

表格 1 能力域与成熟度等级对应关系

能力域	*L1*	*L2*	*L3*	*L4*	*L5*
产品基础能力	√	√	√	√	√
模型工程能力	√	√	√	√	√
业务融合能力	√	√	√	√	√
安全治理能力	√	√	√	√	√
组织运营能力	√	√	√	√	√
持续优化能力	√	√	√	√	√

5.4.3 能力要求

各能力域在不同成熟度等级下的具体要求应符合第6章规定。

成熟度等级评定方法应符合第7章规定。

6 成熟度要求

6.1 产品基础能力

6.1.1 概述

产品基础能力用于评价数字安全大模型产品或系统所具备的基础技术能力、专业能力和安全能力水平。

产品基础能力是数字安全大模型应用能力体系的重要基础，反映模型在知识理解、任务处理、性能表现、资源利用和安全可信等方面的综合能力。

产品基础能力包括：

a）模型能力基础；

b）推理性能；

c）能效水平；

d）安全专业能力；

e）模型安全能力。

6.1.2 模型能力基础

6.1.2.1 概述

模型能力基础用于评价数字安全大模型在知识承载、任务理解、逻辑推理和场景适配等方面所具备的基础能力水平。

6.1.2.2 评估内容

模型能力基础评估包括：

a）模型知识能力；

b）模型推理能力；

c）模型泛化能力；

d）场景适配能力。

6.1.2.3 评估要点

评估应重点关注：

a）模型架构先进性；

b）模型知识覆盖范围；

c）复杂任务理解与处理能力；

d）跨场景迁移能力；

e）数字安全领域适配能力。

6.1.2.4 评估依据

包括但不限于：

a）模型技术文档；

b）模型能力测试报告；

c）模型评测结果；

d）业务应用验证材料；

e）第三方评测报告。

6.1.2.5 能力特征

模型能够支持数字安全领域基础知识理解和任务处理，并具备一定的推理和泛化能力。

模型能力与业务需求保持适配，并能够根据场景需求持续优化。

6.1.2.6 成熟度要求

L1：具备基础模型能力，能够完成通用问答和简单任务处理。
L2：具备数字安全领域基础知识理解能力，能够支持部分安全场景应用。
L3：建立模型选型与能力管理机制，实现模型能力与业务场景匹配。
L4：建立模型能力评价体系，能够持续优化模型能力结构。
L5：建立模型持续演进机制，实现模型能力与业务需求动态协同优化。

6.1.3 推理性能

6.1.3.1 概述

推理性能用于评价数字安全大模型在响应效率、处理能力和服务稳定性等方面的能力水平。

6.1.3.2 评估内容

推理性能评估包括：
a）响应效率；
b）吞吐能力；
c）并发处理能力；
d）服务稳定性。

6.1.3.3 评估要点

评估应重点关注：
a）请求响应时效性；
b）推理处理效率；
c）高并发场景支撑能力；
d）服务连续性；
e）性能优化机制。

6.1.3.4 评估依据

包括但不限于：
a）性能测试报告；
b）系统运行监测记录；
c）性能评测结果；
d）运维管理记录；
e）第三方评测报告。

6.1.3.5 能力特征

模型能够稳定提供推理服务，并满足数字安全场景对实时性和可靠性的要求。

6.1.3.6 成熟度要求

L1：具备基础推理服务能力。
L2：能够满足特定业务场景性能需求。
L3：建立性能监测和优化机制。

L4：建立性能量化评价体系，实现持续优化。
L5：实现性能动态调优和自适应优化。

6.1.4 能效水平

6.1.4.1 概述

能效水平用于评价数字安全大模型资源利用效率和运行成本控制能力。

6.1.4.2 评估内容

能效水平评估包括：
a）算力利用能力；
b）资源利用效率；
c）运行成本控制能力；
d）资源调度能力。

6.1.4.3 评估要点

评估应重点关注：
a）资源利用率；
b）单位资源产出能力；
c）资源调度效率；
d）运行成本优化能力；
e）绿色低碳能力。

6.1.4.4 评估依据

包括但不限于：
a）资源监测报告；
b）运行日志；
c）成本分析报告；
d）运维管理记录；
e）第三方测试报告。

6.1.4.5 能力特征

模型运行资源利用合理，能够兼顾业务效果与资源消耗，实现成本与效益平衡。

6.1.4.6 成熟度要求

L1：能够完成模型部署和运行。
L2：能够监测资源消耗情况。
L3：建立资源利用评价机制。
L4：建立能效评价体系并开展优化。
L5：实现资源利用持续优化和动态调度。

6.1.5 安全专业能力

6.1.5.1 概述

安全专业能力用于评价数字安全大模型在数字安全领域知识理解、分析推理和任务支撑方面的专业水平。

6.1.5.2 评估内容

安全专业能力评估包括：

a）安全知识理解能力；

b）安全分析能力；

c）安全研判能力；

d）安全任务支撑能力。

6.1.5.3 评估要点

评估应重点关注：

a）漏洞分析能力；

b）威胁分析能力；

c）攻击溯源能力；

d）安全运营支撑能力；

e）安全知识问答能力。

6.1.5.4 评估依据

包括但不限于：

a）专业能力测试结果；

b）业务验证结果；

c）应用案例材料；

d）第三方评测结果；

e）专家评审意见。

6.1.5.5 能力特征

模型能够理解数字安全专业知识，并支持安全分析、研判和运营等典型业务场景。

6.1.5.6 成熟度要求

L1：具备基础安全知识理解能力。

L2：能够支持单类安全业务场景。

L3：能够支持多类安全分析任务。

L4：建立专业能力评测体系并持续优化。

L5：形成持续进化的数字安全知识能力体系。

6.1.6 模型安全能力

6.1.6.1 概述

模型安全能力用于评价数字安全大模型在安全可信、风险防控和安全防护方面的能力水平。

6.1.6.2 评估内容

模型安全能力评估包括：

a）模型风险防护能力；

b）内容安全能力；
c）数据保护能力；
d）可信运行能力。

6.1.6.3 评估要点

评估应重点关注：
a）提示注入防护能力；
b）敏感信息保护能力；
c）越权调用防护能力；
d）对抗攻击防护能力；
e）内容安全保障能力。

6.1.6.4 评估依据

包括但不限于：
a）安全测试报告；
b）红队测试结果；
c）风险评估报告；
d）运行监测记录；
e）第三方安全评测报告。

6.1.6.5 能力特征

模型具备基础安全防护能力，能够有效降低模型运行风险，保障模型可信运行。

6.1.6.6 成熟度要求

L1：具备基础安全防护措施。
L2：能够识别常见模型安全风险。
L3：建立模型安全管理机制。
L4：建立模型安全评估与监测体系。
L5：实现模型安全持续优化和主动防御。

6.2 模型工程能力

6.2.1 概述

模型工程能力用于评价数字安全大模型在部署实施、运行管理、能力扩展、系统集成和工程化应用等方面的能力水平。

模型工程能力是连接模型能力与业务应用的重要桥梁，反映组织将大模型能力转化为实际生产力的工程化支撑能力。

模型工程能力包括：
a）部署与运行能力；
b）知识增强能力；
c）工具调用能力；
d）智能体能力；

e）集成与协同能力。

6.2.2 部署与运行能力

6.2.2.1 概述

部署与运行能力用于评价数字安全大模型在部署实施、资源管理、运行维护和服务保障等方面的能力水平。

6.2.2.2 评估内容

部署与运行能力评估包括：

a）部署实施能力；

b）运行管理能力；

c）资源管理能力；

d）运维保障能力。

6.2.2.3 评估要点

评估应重点关注：

a）多环境部署能力；

b）弹性扩展能力；

c）运行监测能力；

d）运维自动化能力；

e）服务保障能力。

6.2.2.4 评估依据

包括但不限于：

a）系统架构文档；

b）部署方案；

c）运维记录；

d）运行监测报告；

e）第三方测试报告。

6.2.2.5 能力特征

模型能够稳定部署和运行，并具备持续服务保障和资源管理能力。

6.2.2.6 成熟度要求

L1：能够完成基础部署和运行。

L2：具备标准化部署能力。

L3：建立统一运维管理机制。

L4：实现运行状态持续监测和量化管理。

L5：实现自动化运维和智能资源调度。

6.2.3 知识增强能力

6.2.3.1 概述

知识增强能力用于评价数字安全大模型利用外部知识提升专业能力和业务能力的水平。

6.2.3.2 评估内容

知识增强能力评估包括：

a）知识获取能力；

b）知识组织能力；

c）知识检索能力；

d）知识更新能力。

6.2.3.3 评估要点

评估应重点关注：

a）知识库建设能力；

b）知识关联能力；

c）知识检索增强能力；

d）知识实时更新能力；

e）知识质量管理能力。

6.2.3.4 评估依据

包括但不限于：

a）知识库建设文档；

b）知识管理制度；

c）知识更新记录；

d）知识应用案例；

e）评测结果。

6.2.3.5 能力特征

模型能够结合外部知识开展分析和推理，并保持知识的时效性和准确性。

6.2.3.6 成熟度要求

L1：能够使用基础知识资源。

L2：建立知识库并开展知识增强应用。

L3：形成知识管理机制。

L4：建立知识质量评价体系。

L5：实现知识动态更新和持续优化。

6.2.4 工具调用能力

6.2.4.1 概述

工具调用能力用于评价数字安全大模型调用外部工具、系统和服务完成复杂任务的能力水平。

6.2.4.2 评估内容

工具调用能力评估包括：

a）工具接入能力；

b）工具调用能力；

c）任务编排能力；

d）结果处理能力。

6.2.4.3 评估要点

评估应重点关注：

a）工具接入范围；

b）调用稳定性；

c）任务执行能力；

d）流程编排能力；

e）结果反馈能力。

6.2.4.4 评估依据

包括但不限于：

a）系统接口文档；

b）工具接入记录；

c）调用日志；

d）业务验证结果；

e）测试报告。

6.2.4.5 能力特征

模型能够自主调用外部工具完成复杂任务，并形成标准化任务处理流程。

6.2.4.6 成熟度要求

L1：具备基础工具调用能力。

L2：支持单工具调用。

L3：支持多工具协同调用。

L4：建立任务编排和流程管理机制。

L5：实现复杂任务自动编排和执行优化。

6.2.5 智能体能力

6.2.5.1 概述

智能体能力用于评价数字安全大模型在任务规划、自主执行、协同决策和持续学习等方面的能力水平。

6.2.5.2 评估内容

智能体能力评估包括：

a）任务理解能力；

b）任务规划能力；

c）自主执行能力；

d）协同决策能力。

6.2.5.3 评估要点

评估应重点关注：

a）目标理解能力；

b）任务分解能力；

c）执行控制能力；

d）状态感知能力；

e）协同决策能力。

6.2.5.4 评估依据

包括但不限于：

a）智能体设计文档；

b）运行记录；

c）任务执行结果；

d）测试报告；

e）应用案例材料。

6.2.5.5 能力特征

模型能够根据目标自主规划任务并完成执行，在复杂场景下具备一定自主决策能力。

6.2.5.6 成熟度要求

L1：能够完成简单任务响应。

L2：具备基础任务规划能力。

L3：能够自主完成多步骤任务。

L4：具备复杂任务协同执行能力。

L5：实现智能体持续优化和自主协同。

6.2.6 集成与协同能力

6.2.6.1 概述

集成与协同能力用于评价数字安全大模型与业务系统、数据平台以及其他智能体协同工作的能力水平。

6.2.6.2 评估内容

集成与协同能力评估包括：

a）系统集成能力；

b）数据协同能力；

c）模型协同能力；

d）智能体协同能力。

6.2.6.3 评估要点

评估应重点关注：

a）系统互联互通能力；

b）数据共享能力；

c）跨模型协同能力；

d）跨智能体协同能力；

e）业务流程协同能力。

6.2.6.4 评估依据

包括但不限于：

a）系统集成方案；

b）接口规范；

c）运行记录；

d）协同任务案例；

e）测试验证结果。

6.2.6.5 能力特征

模型能够与多类系统和智能体协同运行，实现跨系统、跨场景和跨任务协同。

6.2.6.6 成熟度要求

L1：能够完成基础系统集成。

L2：支持多系统连接。

L3：建立统一集成管理机制。

L4：实现跨系统协同运行。

L5：实现多模型、多智能体协同优化和持续演进。

6.3 业务融合能力

6.3.1 概述

业务融合能力用于评价数字安全大模型与数字安全业务场景融合应用的深度和广度。

业务融合能力反映组织将大模型能力转化为实际业务能力和运营能力的水平，是衡量数字安全大模型应用价值的重要维度。

业务融合能力包括：

a）安全运营融合能力；
b）风险分析融合能力；
c）漏洞管理融合能力；
d）事件响应融合能力；
e）智能协同运营能力。

6.3.2 安全运营融合能力

6.3.2.1 概述

安全运营融合能力用于评价数字安全大模型与安全运营流程融合应用的能力水平。

6.3.2.2 评估内容

安全运营融合能力评估包括：
a）运营流程融合能力；
b）告警分析能力；
c）运营支撑能力；
d）运营自动化能力。

6.3.2.3 评估要点

评估应重点关注：
a）安全运营场景覆盖范围；
b）运营流程嵌入程度；
c）告警分析支撑能力；
d）运营效率提升效果；
e）自动化运营能力。

6.3.2.4 评估依据

包括但不限于：
a）运营流程文档；
b）运营平台运行记录；
c）应用案例材料；
d）业务评估报告；
e）用户反馈材料。

6.3.2.5 能力特征

数字安全大模型能够参与安全运营流程，辅助开展告警分析、事件研判和运营管理工作。

6.3.2.6 成熟度要求

L1：在个别运营场景开展试点应用。
L2：能够辅助完成部分运营任务。
L3：融入安全运营流程并形成标准应用模式。
L4：实现运营过程量化评估和持续优化。
L5：实现智能化安全运营和持续协同优化。

6.3.3 风险分析融合能力

6.3.3.1 概述

风险分析融合能力用于评价数字安全大模型在风险识别、风险分析和风险研判等业务场景中的应用能力。

6.3.3.2 评估内容

风险分析融合能力评估包括：
a）风险识别能力；

b）风险分析能力；

c）风险研判能力；

d）风险评估能力。

6.3.3.3 评估要点

评估应重点关注：

a）风险发现能力；

b）关联分析能力；

c）研判辅助能力；

d）风险评估能力；

e）分析结果可解释性。

6.3.3.4 评估依据

包括但不限于：

a）分析报告；

b）研判结果；

c）业务案例；

d）运行记录；

e）测试验证材料。

6.3.3.5 能力特征

数字安全大模型能够结合多源信息开展风险分析和辅助决策。

6.3.3.6 成熟度要求

L1：支持简单风险分析。

L2：能够支持单类风险分析场景。

L3：能够支持多源数据关联分析。

L4：建立风险分析评价体系。

L5：实现智能风险发现和主动研判。

6.3.4 漏洞管理融合能力

6.3.4.1 概述

漏洞管理融合能力用于评价数字安全大模型在漏洞发现、漏洞分析和漏洞管理等业务场景中的应用能力。

6.3.4.2 评估内容

漏洞管理融合能力评估包括：

a）漏洞分析能力；

b）漏洞处置支撑能力；

c）漏洞知识管理能力；

d）漏洞运营能力。

6.3.4.3 评估要点

评估应重点关注：

a）漏洞理解能力；

b）漏洞关联分析能力；

c）修复建议生成能力；

d）知识积累能力；

e）漏洞闭环管理能力。

6.3.4.4 评估依据

包括但不限于：

a）漏洞分析报告；

b）漏洞管理记录；

c）知识库材料；

d）业务验证结果；

e）测试评估结果。

6.3.4.5 能力特征

数字安全大模型能够支持漏洞全生命周期管理并提升漏洞治理效率。

6.3.4.6 成熟度要求

L1：支持基础漏洞查询和分析。

L2：能够辅助开展漏洞研判。

L3：融入漏洞管理流程。

L4：实现漏洞管理量化分析和持续优化。

L5：实现漏洞治理智能化运营。

6.3.5 事件响应融合能力

6.3.5.1 概述

事件响应融合能力用于评价数字安全大模型在安全事件发现、分析、处置和复盘过程中的应用能力。

6.3.5.2 评估内容

事件响应融合能力评估包括：

a）事件分析能力；

b）事件处置支撑能力；

c）事件复盘能力；

d）响应协同能力。

6.3.5.3 评估要点

评估应重点关注：

a）事件分析效率；

b）处置辅助能力；

c）处置流程协同能力；

d）复盘总结能力；

e）知识沉淀能力。

6.3.5.4 评估依据

包括但不限于：

a）事件处置记录；

b）应急响应报告；

c）复盘报告；

d）协同记录；

e）测试结果。

6.3.5.5 能力特征

数字安全大模型能够支持安全事件全流程处置并提升响应效率。

6.3.5.6 成熟度要求

L1：支持事件信息整理。

L2：能够辅助完成事件分析。

L3：融入事件响应流程。

L4：实现响应过程量化管理。

L5：实现智能化事件响应和持续优化。

6.3.6 智能协同运营能力

6.3.6.1 概述

智能协同运营能力用于评价数字安全大模型在人机协同、多角色协同和多智能体协同运营中的应用能力。

6.3.6.2 评估内容

智能协同运营能力评估包括：

a）人机协同能力；

b）团队协同能力；

c）多智能体协同能力；

d）业务协同能力。

6.3.6.3 评估要点

评估应重点关注：

a）协同机制建设情况；

b）任务协同能力；

c）决策协同能力；

d）跨角色协同能力；

e）协同运营效果。

6.3.6.4 评估依据

包括但不限于：

a）协同机制文件；

b）运行记录；

c）业务案例；

d）评价报告；

e）专家评审意见。

6.3.6.5 能力特征

数字安全大模型能够作为数字员工或智能体参与业务协同，实现人机协同和智能协同运营。

6.3.6.6 成熟度要求

L1：支持简单协同任务。

L2：能够参与部分业务协同。

L3：建立标准化协同机制。

L4：实现跨系统、跨角色协同运营。

L5：实现多智能体协同和智能运营闭环。

6.4 安全治理能力

6.4.1 概述

安全治理能力用于评价组织在数字安全大模型应用过程中建立风险管理、安全控制、合规治理和可信运营体系的能力水平。

安全治理能力贯穿数字安全大模型全生命周期，是保障模型安全、业务安全和组织安全的重要基础能力。

安全治理能力包括：

a）数据安全治理能力；

b）模型安全治理能力；

c）智能体安全治理能力；

d）合规与伦理治理能力；

e）安全运营治理能力。

6.4.2 数据安全治理能力

6.4.2.1 概述

数据安全治理能力用于评价组织在数字安全大模型应用过程中对数据资源进行安全管理和风险控制的能力水平。

6.4.2.2 评估内容

数据安全治理能力评估包括：

a）数据分类分级能力；

b）数据安全保护能力；

c）数据流转管理能力；

d）数据风险控制能力。

6.4.2.3 评估要点

评估应重点关注：

a）数据分类分级机制；

b）敏感数据识别能力；

c）数据访问控制机制；

d）数据脱敏和匿名化能力；

e）数据流转审计能力。

6.4.2.4 评估依据

包括但不限于：

a）数据管理制度；

b）数据安全策略；

c）审计记录；

d）风险评估报告；

e）第三方评估报告。

6.4.2.5 能力特征

组织建立覆盖数据采集、存储、处理、共享和销毁全过程的数据安全治理体系。

6.4.2.6 成熟度要求

L1：具备基础数据管理措施。

L2：建立数据安全管理制度。

L3：建立数据分类分级与访问控制机制。

L4：建立数据安全监测和风险评估体系。

L5：实现数据安全全生命周期动态治理。

6.4.3 模型安全治理能力

6.4.3.1 概述

模型安全治理能力用于评价组织识别、控制和管理大模型安全风险的能力水平。

6.4.3.2 评估内容

模型安全治理能力评估包括：

a）模型风险识别能力；

b）模型安全控制能力；

c）模型可信管理能力；

d）模型风险处置能力。

6.4.3.3 评估要点

评估应重点关注：

a）提示注入防护能力；

b）越权调用防护能力；

c）对抗攻击防护能力；

d）模型幻觉控制能力；

e）模型可信性管理能力。

6.4.3.4 评估依据

包括但不限于：

a）模型安全测试报告；

b）红队测试记录；

c）风险评估报告；

d）运行监测记录；

e）整改闭环记录。

6.4.3.5 能力特征

组织建立模型安全管理体系，能够持续发现、评估和控制模型安全风险。

6.4.3.6 成熟度要求

L1：具备基础模型安全防护措施。

L2：建立模型安全风险识别机制。

L3：建立模型安全管理体系。

L4：建立模型安全量化评估体系。

L5：实现模型安全主动防御和持续优化。

6.4.4 智能体安全治理能力

6.4.4.1 概述

智能体安全治理能力用于评价组织对数字员工、智能体以及多智能体系统进行安全管理和风险控制的能力水平。

6.4.4.2 评估内容

智能体安全治理能力评估包括：

a）智能体身份管理能力；

b）智能体权限管理能力；

c）智能体行为控制能力；

d）智能体责任追溯能力。

6.4.4.3 评估要点

评估应重点关注：

a）智能体身份标识机制；

b）权限分配与控制机制；

c）行为监测与审计机制；

d）工具调用约束机制；

e）责任追溯机制。

6.4.4.4 评估依据

包括但不限于：

a）身份管理制度；

b）权限管理策略；

c）行为审计日志；

d）安全测试记录；

e）风险处置记录。

6.4.4.5 能力特征

组织建立数字员工和智能体安全治理体系，实现身份可识别、行为可审计、责任可追溯。

6.4.4.6 成熟度要求

L1：具备基础智能体管理措施。

L2：建立智能体身份管理机制。

L3：建立智能体权限管理体系。

L4：建立智能体行为监测和风险控制体系。

L5：实现智能体全生命周期安全治理。

6.4.5 合规与伦理治理能力

6.4.5.1 概述

合规与伦理治理能力用于评价组织保障数字安全大模型应用符合法律法规、行业规范和伦理要求的能力水平。

6.4.5.2 评估内容

合规与伦理治理能力评估包括：

a）法律合规管理能力；

b）伦理治理能力；

c）风险评估能力；

d）审查管理能力。

6.4.5.3 评估要点

评估应重点关注：

a）法律法规符合性；

b）伦理规范建设情况；

c）风险评估机制；

d）内容安全审查机制；

e）责任管理机制。

6.4.5.4 评估依据

包括但不限于：

a）制度文件；

b）风险评估报告；

c）审查记录；

d）合规审计报告；

e）整改记录。

6.4.5.5 能力特征

组织建立完善的合规与伦理治理机制，能够持续识别和控制合规风险。

6.4.5.6 成熟度要求

L1：关注法律合规要求。

L2：建立基础合规管理制度。

L3：建立合规与伦理治理体系。

L4：建立风险评估与持续监测机制。

L5：实现合规治理持续优化和主动治理。

6.4.6 安全运营治理能力

6.4.6.1 概述

安全运营治理能力用于评价组织建立数字安全大模型安全运营、风险监测和持续改进机制的能力水平。

6.4.6.2 评估内容

安全运营治理能力评估包括：

a）风险监测能力；

b）安全运营能力；

c）事件处置能力；

d）持续改进能力。

6.4.6.3 评估要点

评估应重点关注：

a）风险发现能力；

b）安全告警能力；

c）事件响应能力；

d）运营分析能力；

e）持续改进能力。

6.4.6.4 评估依据

包括但不限于：

a）监测记录；

b）运营报告；

c）事件处置记录；

d）改进记录；

e）评估报告。

6.4.6.5 能力特征

组织建立覆盖监测、分析、响应和改进全过程的安全运营治理体系。

6.4.6.6 成熟度要求

L1：具备基础安全运营能力。

L2：建立安全运营管理机制。

L3：建立安全运营治理体系。

L4：建立量化运营评价机制。

L5：实现安全运营持续优化和主动治理。

6.5 组织运营能力

6.5.1 概述

组织运营能力用于评价组织在数字安全大模型应用过程中建立管理体系、运营体系和保障体系的能力水平。

组织运营能力反映组织将数字安全大模型、数字员工和智能体纳入组织管理体系，实现战略规划、组织协同、资源配置和持续运营的能力。

组织运营能力包括：

a）AI战略与顶层设计能力；

b）数字员工身份管理能力；

c）数字员工组织管理能力；
d）人机协同运营能力；
e）人才与文化建设能力。

6.5.2 AI战略与顶层设计能力

6.5.2.1 概述

AI战略与顶层设计能力用于评价组织规划、推动和管理数字安全大模型应用的能力水平。

6.5.2.2 评估内容

AI战略与顶层设计能力评估包括：
a）战略规划能力；
b）发展路线设计能力；
c）资源统筹能力；
d）组织推动能力。

6.5.2.3 评估要点

评估应重点关注：
a）AI战略规划情况；
b）发展目标明确性；
c）资源保障机制；
d）组织推动机制；
e）长期发展规划。

6.5.2.4 评估依据

包括但不限于：
a）战略规划文件；
b）年度建设方案；
c）组织管理制度；
d）资源投入记录；
e）评估报告。

6.5.2.5 能力特征

组织建立数字安全大模型发展战略，形成统筹规划、协同推进和持续发展的建设机制。

6.5.2.6 成熟度要求

L1：开始关注数字安全大模型发展方向。
L2：制定数字安全大模型建设规划。
L3：建立组织级战略规划体系。
L4：建立战略实施监测与评估机制。
L5：实现战略持续优化和创新引领。

6.5.3 数字员工身份管理能力

6.5.3.1 概述

数字员工身份管理能力用于评价组织建立数字员工和智能体身份管理体系的能力水平。

6.5.3.2 评估内容

数字员工身份管理能力评估包括：
a）身份定义能力；
b）身份管理能力；
c）身份认证能力；
d）身份审计能力。

6.5.3.3 评估要点

评估应重点关注：

a）身份标识机制；

b）身份认证机制；

c）身份生命周期管理机制；

d）身份审计机制；

e）身份责任管理机制。

6.5.3.4 评估依据

包括但不限于：

a）身份管理制度；

b）数字员工档案；

c）认证记录；

d）审计记录；

e）运行记录。

6.5.3.5 能力特征

组织建立统一的数字员工身份管理体系，实现身份可识别、职责可定义、行为可追溯。

6.5.3.6 成熟度要求

L1：建立基础数字员工身份信息。

L2：建立数字员工身份管理机制。

L3：建立数字员工身份生命周期管理体系。

L4：实现身份动态管理和审计监督。

L5：实现数字员工身份体系持续优化和智能管理。

6.5.4 数字员工组织管理能力

6.5.4.1 概述

数字员工组织管理能力用于评价组织建立数字员工岗位体系、职责体系和组织协同体系的能力水平。

6.5.4.2 评估内容

数字员工组织管理能力评估包括：

a）岗位定义能力；

b）职责管理能力；

c）任务管理能力；

d）组织协同能力。

6.5.4.3 评估要点

评估应重点关注：

a）岗位映射机制；

b）职责分工机制；

c）任务分配机制；

d）组织协同机制；

e）绩效管理机制。

6.5.4.4 评估依据

包括但不限于：

a）岗位体系文件；

b）职责说明文件；

c）任务管理记录；
d）协同运行记录；
e）绩效评价材料。

6.5.4.5 能力特征

组织建立数字员工组织体系，实现岗位明确、职责清晰、协同有序。

6.5.4.6 成熟度要求

L1：数字员工参与个别业务活动。
L2：建立数字员工岗位管理机制。
L3：建立数字员工组织管理体系。
L4：实现数字员工组织运行量化管理。
L5：实现数字员工组织体系持续优化和动态调整。

6.5.5 人机协同运营能力

6.5.5.1 概述

人机协同运营能力用于评价组织建立人类员工与数字员工协同工作机制的能力水平。

6.5.5.2 评估内容

人机协同运营能力评估包括：
a）协同机制建设能力；
b）任务协同能力；
c）决策协同能力；
d）运营协同能力。

6.5.5.3 评估要点

评估应重点关注：
a）人机协同流程；
b）人工接管机制；
c）协同决策机制；
d）责任划分机制；
e）协同效率提升情况。

6.5.5.4 评估依据

包括但不限于：
a）协同管理制度；
b）业务流程文件；
c）运行记录；
d）案例材料；
e）评估报告。

6.5.5.5 能力特征

组织建立稳定的人机协同机制，实现人类员工与数字员工协同开展业务活动。

6.5.5.6 成熟度要求

L1：数字员工辅助开展个别任务。
L2：建立基础人机协同机制。
L3：形成标准化人机协同流程。
L4：实现人机协同量化运营。
L5：实现人机协同持续优化和智能协同决策。

6.5.6 人才与文化建设能力

6.5.6.1 概述

人才与文化建设能力用于评价组织培养数字安全大模型应用人才和建设创新文化的能力水平。

6.5.6.2 评估内容

人才与文化建设能力评估包括：

a）人才培养能力；

b）能力建设能力；

c）知识传播能力；

d）创新文化建设能力。

6.5.6.3 评估要点

评估应重点关注：

a）人才培养机制；

b）培训体系建设情况；

c）能力认证机制；

d）知识共享机制；

e）创新文化建设情况。

6.5.6.4 评估依据

包括但不限于：

a）培训计划；

b）培训记录；

c）能力认证材料；

d）知识管理材料；

e）文化建设材料。

6.5.6.5 能力特征

组织形成数字安全大模型人才培养体系和创新文化体系，能够持续支撑能力建设和创新发展。

6.5.6.6 成熟度要求

L1：开展基础培训和能力建设活动。

L2：建立人才培养机制。

L3：建立组织级人才发展体系。

L4：建立人才能力评价和持续提升机制。

L5：形成持续创新的人才生态和组织文化。

6.6 持续优化能力

6.6.1 概述

持续优化能力用于评价组织对数字安全大模型、数字员工和智能体开展持续评估、持续运营、持续改进和持续创新的能力水平。

持续优化能力反映组织推动数字安全大模型应用不断提升业务价值、运营效率和创新能力的能力，是实现高成熟度数字安全大模型应用的重要保障。

持续优化能力包括：

a）能力评价体系建设能力；

b）能力运营管理能力；

c）知识运营能力；

d）模型持续进化能力；

e）创新生态建设能力。

6.6.2 能力评价体系建设能力

6.6.2.1 概述

能力评价体系建设能力用于评价组织建立数字安全大模型、数字员工和智能体能力评价机制的能力水平。

6.6.2.2 评估内容

能力评价体系建设能力评估包括：

a）评价指标建设能力；

b）评价机制建设能力；

c）评价实施能力；

d）评价改进能力。

6.6.2.3 评估要点

评估应重点关注：

a）评价指标体系建设情况；

b）业务价值评价机制；

c）运营效果评价机制；

d）能力评估机制；

e）评价结果应用机制。

6.6.2.4 评估依据

包括但不限于：

a）评价体系文件；

b）评价报告；

c）绩效分析报告；

d）改进记录；

e）运营分析材料。

6.6.2.5 能力特征

组织建立统一的能力评价体系，能够持续评估数字安全大模型应用成效和业务价值。

6.6.2.6 成熟度要求

L1：开展基础能力评价活动。

L2：建立能力评价机制。

L3：形成统一评价体系。

L4：实现评价指标量化管理。

L5：实现评价驱动的持续优化闭环。

6.6.3 能力运营管理能力

6.6.3.1 概述

能力运营管理能力用于评价组织对数字安全大模型、数字员工和智能体开展持续运营和优化管理的能力水平。

6.6.3.2 评估内容

能力运营管理能力评估包括：

a）能力运营能力；

b）服务运营能力；

c）运营分析能力；

d）运营优化能力。

6.6.3.3 评估要点

评估应重点关注：

a）能力运营机制建设情况；
b）运营监测机制；
c）运营分析机制；
d）运营优化机制；
e）运营价值提升情况。

6.6.3.4 评估依据

包括但不限于：
a）运营制度文件；
b）运营报告；
c）运行记录；
d）分析报告；
e）改进记录。

6.6.3.5 能力特征

组织建立数字安全大模型能力运营体系，实现持续监测、持续分析和持续优化。

6.6.3.6 成熟度要求

L1：开展基础运营活动。
L2：建立运营管理机制。
L3：建立组织级能力运营体系。
L4：实现运营数据量化分析。
L5：实现能力运营持续优化和价值提升。

6.6.4 知识运营能力

6.6.4.1 概述

知识运营能力用于评价组织对数字安全知识资产开展建设、管理、更新和运营的能力水平。

6.6.4.2 评估内容

知识运营能力评估包括：
a）知识建设能力；
b）知识管理能力；
c）知识更新能力；
d）知识应用能力。

6.6.4.3 评估要点

评估应重点关注：
a）知识库建设情况；
b）知识更新机制；
c）知识质量管理机制；
d）知识共享机制；
e）知识应用效果。

6.6.4.4 评估依据

包括但不限于：
a）知识管理制度；
b）知识库材料；
c）更新记录；
d）应用案例；
e）评估报告。

6.6.4.5 能力特征

组织建立知识运营体系，实现知识持续积累、持续更新和持续应用。

6.6.4.6 成熟度要求

L1：建立基础知识资源。

L2：建立知识管理机制。

L3：建立知识运营体系。

L4：实现知识质量量化管理。

L5：实现知识持续进化和智能运营。

6.6.5 模型持续进化能力

6.6.5.1 概述

模型持续进化能力用于评价组织持续提升数字安全大模型能力水平和应用效果的能力。

6.6.5.2 评估内容

模型持续进化能力评估包括：

a）模型优化能力；

b）模型迭代能力；

c）能力增强能力；

d）效果提升能力。

6.6.5.3 评估要点

评估应重点关注：

a）模型优化机制；

b）模型迭代机制；

c）反馈闭环机制；

d）能力增强机制；

e）持续改进机制。

6.6.5.4 评估依据

包括但不限于：

a）模型迭代记录；

b）优化方案；

c）评测报告；

d）反馈记录；

e）改进报告。

6.6.5.5 能力特征

组织建立模型持续优化机制，能够根据业务需求和运行反馈持续提升模型能力。

6.6.5.6 成熟度要求

L1：开展基础模型优化工作。

L2：建立模型优化机制。

L3：建立模型迭代管理体系。

L4：实现模型能力量化提升。

L5：实现模型持续进化和自主优化。

6.6.6 创新生态建设能力

6.6.6.1 概述

创新生态建设能力用于评价组织推动数字安全大模型创新应用、生态协同和持续创新发展的能力水平。

6.6.6.2 评估内容

创新生态建设能力评估包括：

a）创新机制建设能力；

b）生态协同能力；

c）创新应用能力；

d）创新推广能力。

6.6.3.3 评估要点

评估应重点关注：

a）创新管理机制；

b）创新项目建设情况；

c）生态合作机制；

d）创新成果转化能力；

e）创新推广能力。

6.6.6.4 评估依据

包括但不限于：

a）创新管理制度；

b）创新项目材料；

c）合作协议；

d）成果转化材料；

e）评估报告。

6.6.6.5 能力特征

组织建立持续创新机制，形成开放协同、持续演进的数字安全大模型创新生态。

6.6.6.6 成熟度要求

L1：开展基础创新探索。

L2：建立创新管理机制。

L3：建立创新项目管理体系。

L4：形成创新评价和推广机制。

L5：形成持续创新和生态协同发展能力。

7 评估方法

7.1 评估原则

7.1.1 客观性原则

评估应基于事实证据、运行数据、测试结果和支撑材料开展，不应依赖主观判断作为唯一依据。

7.1.2 系统性原则

评估应覆盖产品基础能力、模型工程能力、业务融合能力、安全治理能力、组织运营能力和持续优化能力等全部能力域。

7.1.3 一致性原则

评估过程、评估方法和评分规则应保持一致，确保评估结果具有可重复性和可比性。

7.1.4 持续改进原则

评估应关注组织能力建设和持续优化情况，鼓励组织持续提升数字安全大模型应用能力。

7.2 评估对象分类

7.2.1 产品型评估

产品型评估适用于数字安全大模型产品、数字安全智能体产品以及相关平台产品。

产品型评估重点关注产品基础能力、模型工程能力和模型安全能力。

产品型评估结果记为：DSL²M²-P（Product）。

7.2.2 组织型评估

组织型评估适用于政府部门、企事业单位、安全运营机构以及其他数字安全大模型应用组织。

组织型评估重点关注业务融合能力、安全治理能力、组织运营能力和持续优化能力。

组织型评估结果记为：DSL²M²-O（Organization）。

7.3 评分方法

7.3.1 评分结构

DSL²M²采用三级评分结构。

a）能力域评分；

b）能力子域评分；

c）综合评分。

7.3.2 能力子域评分

每个能力子域应依据第6章规定的：

a）评估内容；

b）评估要点；

c）评估依据；

d）能力特征；

e）成熟度要求；

开展评分。

7.3.3 评分等级

能力子域评分采用五级评价方式。

表格 2 评分等级

评分等级	*描述*
1	初步具备
2	基本具备
3	较好具备
4	完全具备
5	行业领先

7.3.4 能力域评分

能力域得分由所属能力子域得分加权计算。

能力域满分为100分。

能力域评分应综合考虑：

a）能力建设情况；

b）应用实施情况；

c）运行效果情况；

d）持续优化情况。

7.4 权重计算

7.4.1 产品型评估权重

产品型评估采用表3规定的权重。

表格 3 产品型评估权重

能力域	*权重*
产品基础能力	30%
模型工程能力	25%
业务融合能力	10%
安全治理能力	15%
组织运营能力	5%
持续优化能力	15%

产品型评估综合得分按照公式（1）计算：

$DSL^2M^2\text{-}P = 0.30A + 0.25B + 0.10C + 0.15D + 0.05E + 0.15F$ （1）

式中：

A——产品基础能力得分；

B——模型工程能力得分；

C——业务融合能力得分；

D——安全治理能力得分；

E——组织运营能力得分；

F——持续优化能力得分。

7.4.2 组织型评估权重

组织型评估采用表4规定的权重。

表格 4 组织型评估权重

能力域	*权重*
产品基础能力	10%
模型工程能力	15%
业务融合能力	25%
安全治理能力	20%
组织运营能力	20%
持续优化能力	10%

组织型评估综合得分按照公式（2）计算：

$DSL^2M^2\text{-}O = 0.10A + 0.15B + 0.25C + 0.20D + 0.20E + 0.10F$

式中各参数含义同公式（1）。

7.5 成熟度等级判定

7.5.1 判定原则

成熟度等级判定应同时满足：

a）综合得分要求；

b）能力域最低要求。

7.5.2 综合得分等级

综合得分与成熟度等级对应关系见表5。

表格 5 综合得分等级划分

综合得分	*成熟度等级*

＜20	L1 初始级
20～39	L2 可管理级
40～59	L3 已定义级
60～79	L4 量化管理级
≥80	L5 优化创新级

7.5.3 能力域门槛要求

成熟度等级除满足综合得分要求外，还应满足表6规定的能力域门槛要求。

表格 6 能力域门槛要求

等级	*门槛要求*
L2	安全治理能力达到L2水平
L3	安全治理能力达到L3水平
L4	安全治理能力和组织运营能力达到L3水平
L5	六个能力域均达到L4水平及以上

7.5.4 降级规则

当组织未满足对应等级能力域门槛要求时，应按照满足条件的最高等级进行认定。

7.6 评估流程

7.6.1 评估准备

评估机构应明确评估对象、评估范围和评估类型。

7.6.2 材料审核

评估机构应对评估对象提交的制度文件、技术文档、测试报告、运行记录及相关证明材料进行审核。

7.6.3 能力评估

评估机构应依据第6章规定开展能力评估。

7.6.4 综合评分

评估机构应依据本章规定计算能力域得分和综合得分。

7.6.5 等级判定

评估机构应依据综合得分和能力域门槛要求确定成熟度等级。

7.6.6 结果确认

评估机构应形成评估报告并确认评估结果。

7.7 评估结果

7.7.1 结果形式

评估结果应包括：

a）评估类型；

b）综合得分；

c）成熟度等级；

d）能力域得分；

e）改进建议。

7.7.2 结果表示

产品型评估结果表示为：DSL^2M^2-P-Lx ，其中：x为成熟度等级。

示例：DSL^2M^2-P-L4 表示产品达到数字安全大模型应用成熟度模型产品型评估四级水平。

组织型评估结果表示为：DSL^2M^2-O-Lx ，其中：x为成熟度等级。

示例：DSL^2M^2-O-L5 表示组织达到数字安全大模型应用成熟度模型组织型评估五级水平。

参考文献

[1] Capability Maturity Model (CMM), Software Engineering Institute (SEI)

[2] Capability Maturity Model Integration (CMMI), CMMI Institute

[3] NIST AI Risk Management Framework (AI RMF)

[4] OWASP Top 10 for Large Language Model Applications

[5] ISO/IEC TR 24028 Information technology — Artificial intelligence — Overview of trustworthiness in artificial intelligence

[6] ISO/IEC TR 24368 Overview of ethical and societal concerns in artificial intelligence

附录A

《产品基础能力评分指南》